

ORIGINALES

Concordancia en la medición de presión arterial entre diferentes profesionales sanitarios. ¿Son fiables los esfigmomanómetros de mercurio?

M. Ripollés Ortí^a, E. Martín Rioboó^b, A. Díaz Moreno^a, B. Aranguren Baena^c, M. Murcia Simón^c, A. Toledano Medina^a y F.J. Fonseca del Pozo^a

Centro de Salud Huerta de la Reina. Córdoba. Consultorio de Pedro Abad. Centro de salud de Montoro (Córdoba).

Objetivos. Valorar fiabilidad en términos de concordancia interobservador de las tomas de presión arterial (PA). Varios profesionales sanitarios y sistemas de medición. Influencia de la experiencia del observador.

Diseño. Estudio observacional, descriptivo transversal.

Emplazamiento. Centro de salud urbano, Córdoba.

Participantes. Un total de 131 pacientes hipertensos, aleatorizados, pertenecientes a una unidad de atención funcional; se excluyó a 11.

Mediciones. Para disminuir variabilidad: curso sobre la toma correcta de la PA, otoscopia y comprobación de agudeza visual de los observadores, calibración y validación de aparatos de medida, tiempo limitado y cegamiento de medidas. Se realizaron 4 mediciones de PA por paciente, tres con esfigmomanómetro de mercurio (EMM) (dos simultáneas, una individual) y una con aparato automático. Se recogieron variables descriptivas, clínicas y somatométricas. La concordancia interobservador se evaluó mediante el coeficiente de correlación intraclass (CCI), método de la media de las diferencias (MMD) e índice de concordancia simple (ICS). Un CCI > de 0,75 se consideró aceptable. Una diferencia > 5 mmHg se consideró relevante clínicamente (MMD e ICS).

Resultados principales. Acuerdo aceptable para el MMD: sólo, PAS y PAD de OBS-1/OBS-2, biauricular, (−6,1/+ 8,9 mmHg) y (−6,8/+ 5,8 mmHg). Resultados más desfavorables: para TAS y TAD: OBS-1/AUTO −20,9/25,0 y −16,4/15,1; OBS-2/AUTO −22,8/24,4 y −16,6/15,2. Resto de intervalos siempre > 10 mmHg; CCI > 0,75 en todas las comparaciones, excepto PAD OBS-1/AUTO y PAD OBS-2/AUTO (0,69 para ambos). Un 41% de comparaciones > 5 mmHg. Sin diferencias en profesional menos experto.

Conclusiones. Inexactitud del método de medida de PA de referencia (EMM) para MMD e ICS.

Conclusiones contradictorias según método de medida. Diferencias no aceptables clínicamente.

Palabras clave: Hipertensión arterial. Fiabilidad. Concordancia interobservador.

AGREEMENT IN THE MEASUREMENT OF BLOOD PRESSURE BETWEEN DIFFERENT HEALTH PROFESSIONALS. ARE MERCURY SPHYGMOMANOMETERS RELIABLE?

Objectives. To assess reliability in terms of inter-observer agreement of blood pressure (BP) readings. Various health professionals and measuring systems. Influence of observer's experience.

Design. Observational, descriptive, cross-sectional study.

Setting. Urban health centre, Córdoba.

Participants. 131 hypertense, randomised patients, belonging to a functional care unit. 11 were excluded.

Measurements. To reduce variability: course on the right way to take blood pressure, otoscope and verification of visual sharpness of observers, calibration and validation of measuring devices, limited time and blinding of measurements. 4 BP measurements per patient: 3 with mercury sphygmomanometer (2 simultaneously, one individual) and one with an automatic device. Descriptive, clinical and somatometric variables were gathered. Inter-observer agreement was evaluated through the intraclass correlation coefficient (ICC), the mean of differences method (MDM) and the simple concordance index (CI). An ICC > 0.75 was thought acceptable. A difference > 5 mmHg was thought clinically relevant (MDM and CI).

Main results. Acceptable consistency for MDM: alone, systolic and diastolic pressure of OBS 1/ OBS 2, bi-auricular, −6.1/+8.9 mmHg and −6.8/+5.8 mmHg. Less favourable results: for systolic and diastolic pressure: OBS 1/AUTO −20.9/25.0 and −16.4/15.1; OBS 2/AUTO −22.8/24.4 and −16.6/15.2. Remaining intervals always > 10 mmHg; CI > 0.75 in all comparisons except diastolic pressure OBS 1/AUTO and diastolic pressure OBS 2/AUTO (0.69 in both cases). 41% of comparisons were > 5 mmHg. No differences in less expert professionals were found.

Conclusions. Inaccuracy of the standard BP measurement method (mercury sphygmomanometer) for MDM and CI. Contradictory conclusions according to method of measurement. Differences not clinically acceptable.

Key words: Hypertension. Reliability. Inter-observer agreement.

(Aten Primaria 2001; 27: 234-243)

^aMédicos de Familia. ^bMiembro del Grupo de Trabajo de HTA de la Sociedad Andaluza de Medicina Familiar y Comunitaria. ^cResidente de Medicina Familiar y Comunitaria.

Correspondencia: Enrique Martín Rioboó.
C/ Antonio Barroso y Castillo, 10, 4.º, 1.ª, 14006 Córdoba.
Correo electrónico: med027294@nacom.es

Manuscrito aceptado para su publicación el 27-XI-2000.

Introducción

Como es sabido, la hipertensión arterial (HTA) es una enfermedad crónica de gran relevancia en nuestro medio y de manejo fundamental en la atención primaria de salud (APS).

El diagnóstico de la HTA se basa en la medición de cifras elevadas de presión arterial (PA) obtenidas con métodos estándar, los cuales han demostrado epidemiológicamente su utilidad clínica y pronóstica.

La técnica de elección continúa siendo la medición efectuada por un médico adiestrado utilizando un esfigmomanómetro de mercurio (EMM) con la técnica de los tonos de Korotkoff¹. Esta técnica no deja de ser una medición indirecta de una variable orgánica continua, sometida a numerosas fuentes de error, no siempre evitables y pocas veces evitadas. Incluso en condiciones ideales, la discrepancia entre mediciones realizadas según los tonos de Korotkoff y mediciones intraarteriales pueden llegar a ser de 25 mmHg².

Como afirma un reciente editorial de la revista JAMA, probablemente ninguno de los componentes de la atención médica tenga mayor importancia para el posterior cuidado del paciente que el establecimiento de un diagnóstico³.

En HTA serán unos valores muy concretos de PA los que nos permitirán clasificar a nuestros pa-

cientes como hipertensos o no, definir el estadio en el que se encuentran, y en función de dichos valores tomar unas decisiones terapéuticas u otras. De ahí la importancia de obtener unas mediciones lo más próximas posible a las reales, y en cualquier caso obtenidas eliminando en lo posible las fuentes evitables de error.

Dichas fuentes de variabilidad en una medición podrían ser debidas al mismo procedimiento (instrumentos de medida, factores ambientales), al observador (metodología de medición, déficit sensitivos) y debidas al sujeto observado⁴. Además sabemos que la PA está expuesta a una gran variabilidad intrínseca que dificulta aún más su medición de forma sistemática.

A pesar de que numerosos estudios han demostrado la posibilidad de medir la validez y exactitud de estos aparatos de medición, con lo que se asegura la capacidad de la prueba para ser reproducible y con validez externa, son menos los que tratan la posible variabilidad interobservador utilizando diferentes sistemas de medida⁵.

En el presente estudio pretendemos analizar si, una vez eliminadas las posibles fuentes de variabilidad, los valores obtenidos mediante una técnica estándar (EMM), pero por diferentes observadores es coincidente, y si dichos valores concuerdan con los obtenidos por un método alternativo de referencia aceptado (aparato automático validado). Se trata, por tanto, de un estudio de fiabilidad, en términos de concordancia interobservador, llevado a cabo sobre una población de hipertensos conocidos.

Los procedimientos más adecuados para expresar la fiabilidad en la medición de variables cuantitativas continuas, como es el caso de la PA, son el método de la media de las diferencias (MMD)⁶⁻⁹ y el coeficiente de correlación intraclass (CCI)^{6,9-12}.

Pacientes y métodos

Tipo de estudio, ámbito y población

Se realizó un estudio descriptivo, transversal, en un centro de salud urbano de la ciudad de Córdoba entre septiembre y noviembre de 1998.

Participaron en el estudio 131 pacientes, seleccionados a partir de un listado de todos los sujetos hipertensos pertenecientes a una unidad de atención funcional (UAF). Se contactó con ellos vía telefónica o bien en consulta, informándoles de la finalidad del estudio y concertando con ellos una cita, previo consentimiento informado. Se les recordaba de forma verbal o por escrito (según selección) las recomendaciones previas a seguir el día de la toma para evitar la variabilidad intrínseca de la PA y para conseguir tomas tensionales en estado basal^{13,14}.

Criterios de exclusión: presentar arritmia cardíaca, crisis hipertensiva en el momento de la medición o un perímetro braquial superior a 35 mm (por problemas técnicos, al no disponer en el momento del estudio de manguito superior a esa medida en el aparato automático de medición). Se excluyó un total de 11 pacientes.

Método

Etapa preliminar. Orientada a disminuir algunas fuentes de variabilidad. Se realizó un seminario de metodología para una correcta toma de PA, dirigido a todo el personal sanitario que iba a participar (3 médicos), siguiendo los criterios de la Association for the Advancement of Medical Instrumentation (AAMI) y de la British Hypertension Society (BHS)^{15,16}. Se realizó, asimismo, un examen otoscópico y valoración de agudeza visual de dicho personal, todo ello encaminado a evitar el sesgo del observador¹⁷. Se procedió también a calibrar los diferentes EMM que iban a utilizarse, frente a un EMM previamente calibrado por personal técnico, comprobando que ninguno estuviese defectuoso y comparándolos posteriormente entre sí. Se utilizó un medidor automático de presión arterial correctamente validado¹⁸. Con ello se pretendió minimizar los sesgos debidos al método de medida.

Recogida de datos. Además de recoger variables descriptivas y clínicas, a cada paciente se le realizaron 4 tomas de PA sistólica (PAS) y diastólica (PAD):

– Dos lecturas simultáneas, con EMM y estetoscopio biauricular, realizadas por 2 médicos con mayor experiencia clínica (OBS-1, OBS-2).

– Una observación individual, con EMM y estetoscopio simple, realizada por otro médico (OBS-3), con menor experiencia clínica.

– Una lectura con aparato automático validado (AAV).

Dichas mediciones se realizaron siguiendo la metodología propuesta por el VI Informe del Joint National Committee (JNC) y las recomendaciones de la Organización Mundial de la Salud (OMS)^{13,14}.

Los aparatos utilizados fueron:

– EMM marca Speider-Keller convenientemente calibrado por personal técnico del Servicio Andaluz de Salud, con manguito de 26 cm de longitud y una anchura de 12 cm (medidas recomendadas por la American Heart Association para la mayoría de los adultos¹⁹).

– Aparato automático de medida de PA marca OMRON HEM-705 CP validado tanto con criterios de la AAMI como de la BHS¹⁸.

Sistemática. Tras un reposo de varios minutos en la sala de espera, el paciente entraba en una de las 2 consultas preparadas para la recogida de los datos (A y B). Se empezó de manera alternante por una u otra consulta, para evitar errores sistemáticos que se podrían cometer al comenzar siempre por el mismo observador.

En la consulta A se recogieron datos referentes al estadio hipertensivo del paciente, fármacos, control tensional y patologías asociadas, anotándolos en una hoja de datos. Tras 5 minutos de reposo, con el paciente sentado, se realizaban 2 tomas simultáneas con el mismo EMM y estetoscopio biauricular (OBS-1, OBS-2), anotando los resultados de manera independiente y en secreto. Transcurridos 3-5 minutos, se tomaba de nuevo la PA, utilizando el aparato automático (AAV), y anotándolo asimismo de manera ciega (los observadores ocultaban sus medidas para no influenciarse mutuamente). Se procuró que el tiempo máximo de espera hasta la siguiente toma fuese inferior a 10 minutos, utilizando consultas muy próximas entre sí y en un horario fuera de la consulta a demanda habitual, para evitar las aglomeraciones propias de un día normal de visitas y, por lo tanto, disminuir la variabilidad intrínseca de la PA.

En la consulta B se recogieron los datos referentes a edad, género, frecuencia cardíaca, índice de masa corporal (IMC) y perímetro braquial. Tras 5 minutos de reposo, se realizaba una toma de PA por otro médico (OBS-3), utilizando un EMM y estetoscopio simple. Todos los pacientes pasaron de forma consecutiva por ambas consultas.

Análisis estadístico

Se realizó un análisis descriptivo de las siguientes variables: edad, género, IMC y perímetro braquial.

Se calculó la media y desviación estándar tanto para la PAS como para la PAD de cada uno de los observadores.

El grado de acuerdo interobservador se valoró en primer lugar por el MMD^{6,7}. Para ello se determinaron las medias de las diferencias de los valores de PA obtenidos por cada par de observadores para cada sujeto. Se halló la media del conjunto de diferencias y su desviación estándar. Se calculó el intervalo de concordancia (media de las diferencias $\pm 1,96 \times DE$), entre cuyos límites están un 95% de los valores de las diferencias encontradas, y se calcularon también los intervalos de confianza de los límites de concordancia⁶. Este método aporta información gráfica del grado de acuerdo entre los diferentes métodos empleados y permite descubrir determinados sesgos (gráficos de Bland y Altman^{4,6}).

En segundo lugar, se utilizó el coeficiente de correlación intraclass (CCI), que se basa en un modelo de análisis de la variancia (ANOVA) con medidas repetidas. Su valor se extiende entre 0 y 1, siendo 1 un perfecto acuerdo y 0 un acuerdo nulo^{4,10,12}. Se interpretó según la clasificación de Fermanian (tabla 4)²⁰. Se ha considerado aceptable un CCI > 0,75 (valor que representa una fiabilidad excelente^{4,10,12,21}).

Se calculó asimismo, como reflejo de la tendencia observada en la clínica, la proporción de valores que diferían en más de 5 mmHg. También se calculó para una diferencia mayor de 10 mmHg. A este valor le denomina Fleiss «índice de concordancia simple» (IC simple)¹⁰.

Resultados

Variables descriptivas

La media de edad fue de $67,14 \pm 9,65$ años; participaron un 33,3%

TABLA 1. Presión arterial media y desviación estándar obtenidas por los diferentes métodos

| | PAS | DE | PAD | DE |
|-------|--------|-------|-------|-------|
| OBS 1 | 156,02 | 19,55 | 82,82 | 10,50 |
| OBS 2 | 154,71 | 19,41 | 83,45 | 13,93 |
| OBS 3 | 155,75 | 18,29 | 84,62 | 9,38 |
| AAV | 153,95 | 18,93 | 83,51 | 9,93 |

PAS: presión arterial sistólica; PAD: presión arterial diastólica, y DE: desviación estándar.

TABLA 2. Análisis de la media de las diferencias entre las observaciones

| | MD | DE | IC | Límites de concordancia | | | | |
|-----------|-----|------|------|-------------------------|-------|---------------|------|-------------|
| | | | | Inferior | | Superior | | |
| OBS1/OBS2 | | | | | | | | |
| | PAS | 1,4 | 3,8 | 15,0 | -6,1 | (-7,3:-4,9) | 8,9 | (7,7:10,0) |
| | PAD | 0,1 | 4,2 | 12,6 | -6,8 | (-8,1:-5,5) | 5,8 | (4,5:7,1) |
| OBS1/AAV | | | | | | | | |
| | PAS | 2,1 | 11,7 | 45,9 | -20,9 | (-24,5:-17,2) | 25,0 | (21,4:28,6) |
| | PAD | -0,7 | 8,0 | 31,5 | -16,4 | (-18,9:-13,9) | 15,1 | (12,6:17,6) |
| OBS1/OBS3 | | | | | | | | |
| | PAS | 0,3 | 8,7 | 34,0 | -16,7 | (-19,4:-14,0) | 17,3 | (14,6:20,0) |
| | PAD | -1,9 | 6,5 | 25,5 | -14,6 | (-16,6:-12,6) | 10,9 | (8,8:12,9) |
| OBS2/AAV | | | | | | | | |
| | PAS | 0,8 | 12,0 | 47,2 | -22,8 | (-26,5:-19,0) | 24,4 | (20,6:28,1) |
| | PAD | -0,7 | 8,1 | 31,8 | -16,6 | (-19,1:-14,1) | 15,2 | (12,6:17,7) |
| OBS2/OBS3 | | | | | | | | |
| | PAS | -1,1 | 8,7 | 34,1 | -18,1 | (-20,8:-15,4) | 16,0 | (13,3:18,7) |
| | PAD | -2,0 | 6,6 | 26,0 | -15,0 | (-17,0:-12,9) | 11,0 | (8,9:13,1) |
| OBS3/AAV | | | | | | | | |
| | PAS | 1,6 | 9,9 | 38,8 | -17,8 | (-20,8:-14,7) | 21,0 | (17,9:24,1) |
| | PAD | 1,2 | 5,9 | 23,2 | -10,5 | (-12,3:-8,6) | 12,8 | (10,9:14,6) |

MD: media de las diferencias; DE: desviación estándar, e IC: intervalo de concordancia.

de varones y un 66,7% de mujeres.

Se obtuvo un IMC de $31,14 \pm 4,704$ y un perímetro braquial de $30,78 \pm 4,12$ mm.

Estudio de la concordancia

Los valores de la PA media obtenidos por cada observador, así como su desviación estándar se exponen en la tabla 1.

La tabla 2 recoge las medias de las diferencias de las respectivas mediciones de PA en cada sujeto, los límites de concordancia y sus correspondientes intervalos de confianza del 95%. En las condiciones en que hemos medido y

sobre nuestra muestra de sujetos, observamos que la media de las diferencias menor corresponde a PAD OBS-1/OBS-2 (0,1 mmHg), que además presentaba el intervalo de concordancia más estrecho (-6,78/+ 5,78 mmHg), seguido por PAS OBS-1/OBS-2 (1,4 mmHg), con un intervalo de concordancia de (-6,10/+ 8,86 mmHg). En las figuras 1 y 2 se muestran los gráficos correspondientes según el MMD. El intervalo de concordancia más amplio corresponde a PAS OBS-1/AAV (45,86 mmHg) y PAS OBS-2/AAV (47,06 mmHg). En las figuras 3 y 4 se recoge su representación gráfica. Alcanzaron

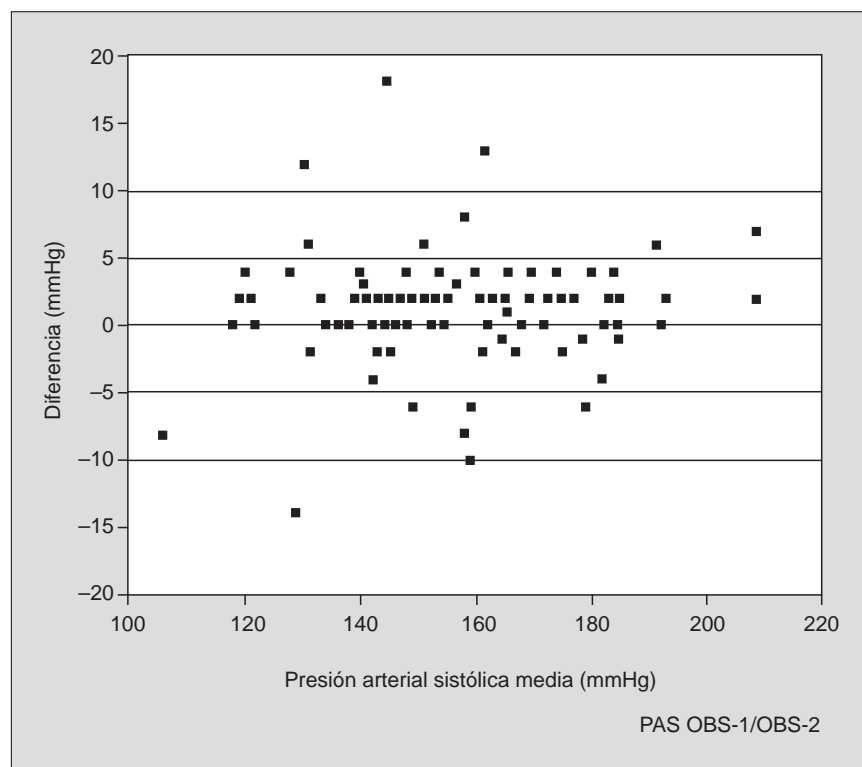


Figura 1. Método de la media de las diferencias individuales (Bland y Altman).

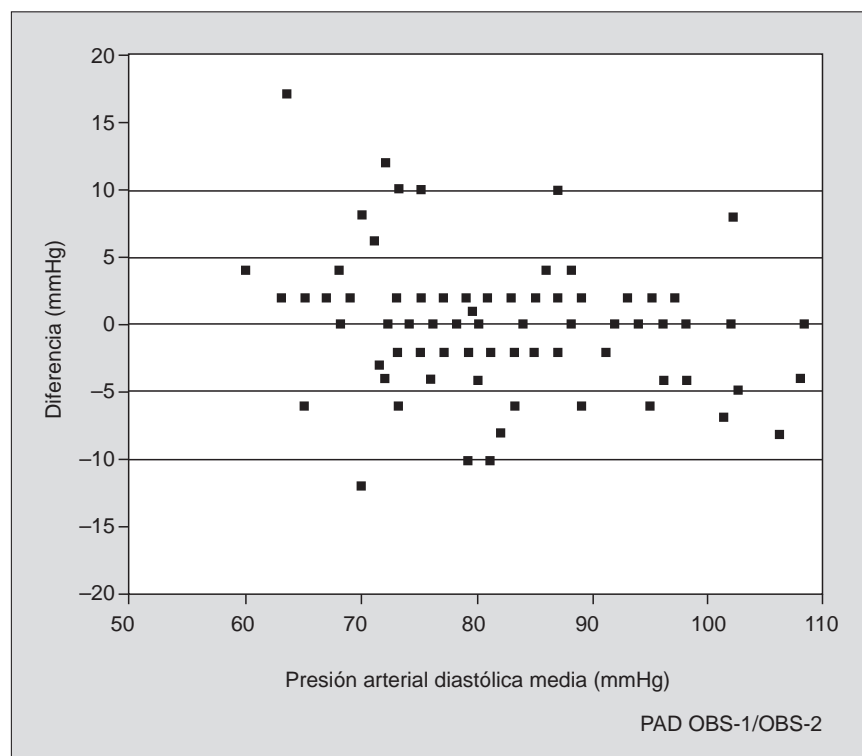


Figura 2. Método de la media de las diferencias individuales (Bland y Altman).

significación estadística las comparaciones correspondientes a PAS OBS-1/OBS-2, PAD OBS-1/OBS-3 y PAD OBS-2/OBS-3.

El CCI y su intervalo de confianza para las diferentes mediciones se exponen en la tabla 3. El CCI mayor corresponde de nuevo a OBS-1/OBS-2, tanto para PAS como para PAD (0,98 y 0,92, respectivamente), considerada como muy buena. El resto de comparaciones se consideran como buenas, excepto la PAD OBS-1/AAV y PAD OBS-2/AAV, que se consideran como moderada, (valores de 0,69 en ambos casos) y son las únicas comparaciones con un CCI $< 0,75$ (figs. 5 y 6).

El IC simple se refleja en la tabla 5. Para OBS-1/OBS-2, un 13,4% de las observaciones de PAS y un 16% de valores de PAD diferían en más de 5 mmHg. La mayor discordancia se observó entre OBS-1/AAV para la PAS (61,7% valores > 5 mmHg) y OBS-2/AAV para la PAS (61,4% valores > 5 mmHg). Del total de comparaciones realizadas, un 41% difería en más de 5 mmHg, y si ampliábamos el intervalo observamos que un 17% difería en más de 10 mmHg.

Discusión

Consideramos que el estudio del grado de fiabilidad en el uso de los distintos aparatos de medición de PA, y con diferentes observadores, es de un evidente interés en la práctica clínica.

Se sabe que hay un cierto grado de error intrínseco en cualquier procedimiento de medición, particularmente cuando el componente principal del proceso es la apreciación subjetiva de un observador. Estos errores pueden llevar a invalidar todo el proceso de atención clínica a un paciente. Aparte de la exactitud de un método (capacidad para valorar aquello que realmente se pretende medir), un atributo muy importante de las mediciones es valorar la fiabilidad o

consistencia de una observación. De hecho, si un procedimiento muestra ausencia de fiabilidad, no es procedente ni siquiera evaluar su exactitud²². Además, es fundamental en un estudio de este tipo neutralizar previamente las fuentes evitables de variabilidad (en este trabajo, etapa preliminar, anotación de datos independiente, condiciones ambientales ideales, etc., todo lo cual ha sido descrito en la sección de método). En el análisis estadístico de nuestros datos, al intentar determinar si en condiciones óptimas de medida todos los observadores obtenían similares resultados, nos encontramos con resultados contradictorios:

1. Al evaluar el grado de acuerdo interobservador por uno de los métodos más aceptados en la bibliografía, el CCI, obteníamos una elevada fiabilidad (definida por CCI > 0,75), excepto para dos de las comparaciones, lo que hacía pensar en una posible intercambiabilidad de los distintos observadores y métodos de medida.
2. En cambio, cuando adoptábamos como criterio de concordancia el MMD (de relevancia complementaria e incluso aún mayor para este tipo de comparaciones)⁶, ya podemos apreciar un buen número de comparaciones que difieren en más de 5 mmHg (prácticamente todas si exceptuamos las obtenidas con el estetoscopio biauricular), llegando incluso a diferencias de 35 mmHg, lo que consideramos clínicamente inaceptable. Ello, incumpliría además de cumplir los criterios de intercambiabilidad propuestos por el National Committee for Clinical Laboratory Standards NCCLS^{4,23}.

3. El IC simple evidencia resultados similares al MMD, con un total del 41% de comparaciones mayores de 5 mmHg, pero a pesar de ser muy gráfico y acercarnos a la realidad clínica, no es valorable estadísticamente, ya que puede

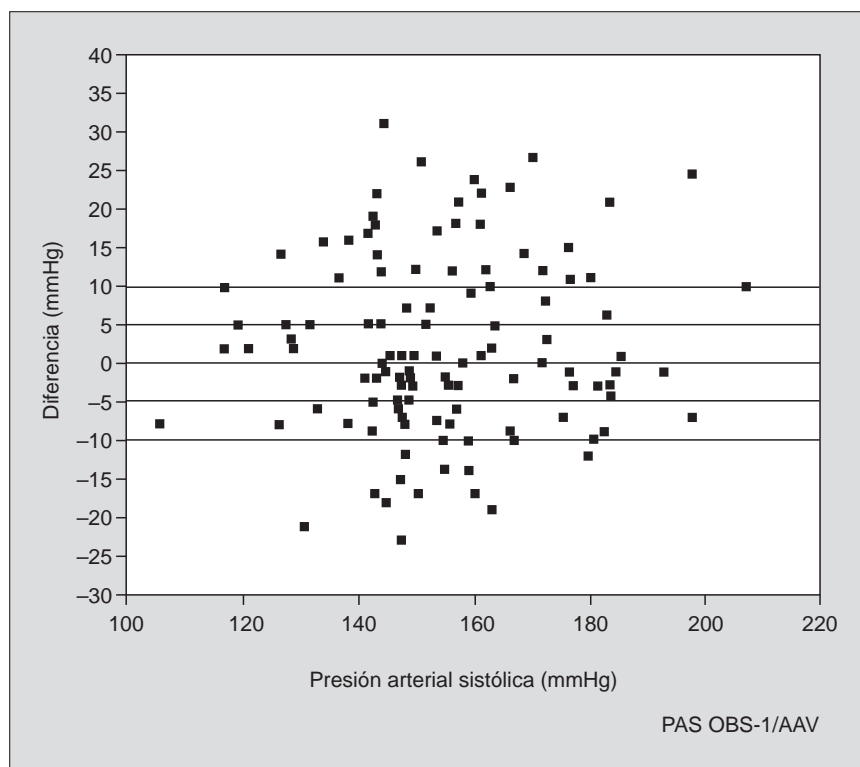


Figura 3. Método de la media de las diferencias individuales (Bland y Altman).

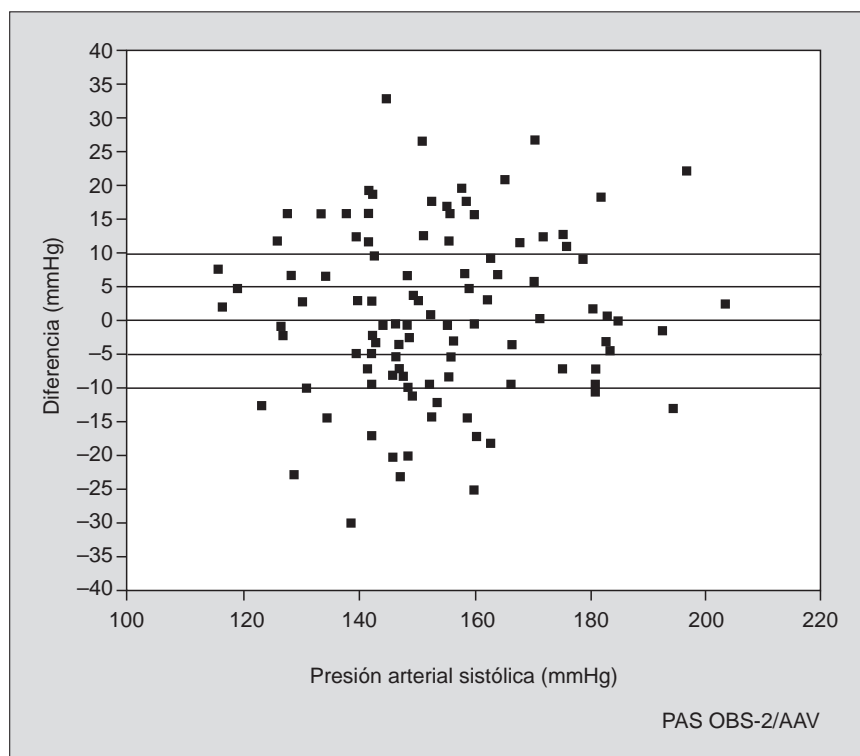


Figura 4. Método de la media de las diferencias individuales (Bland y Altman).

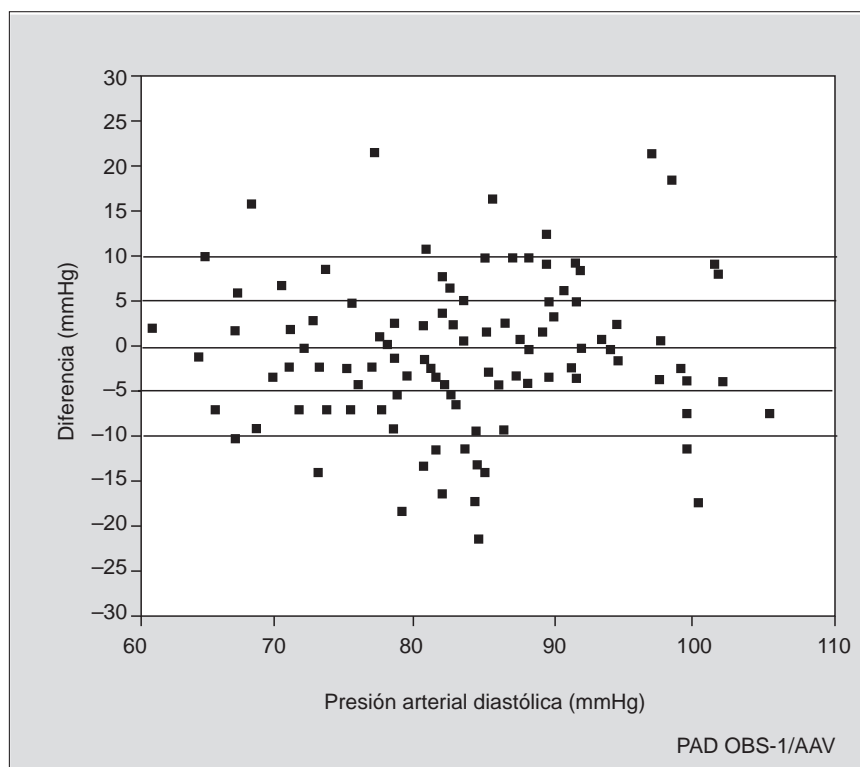


Figura 5. Método de la media de las diferencias individuales (Bland y Altman).

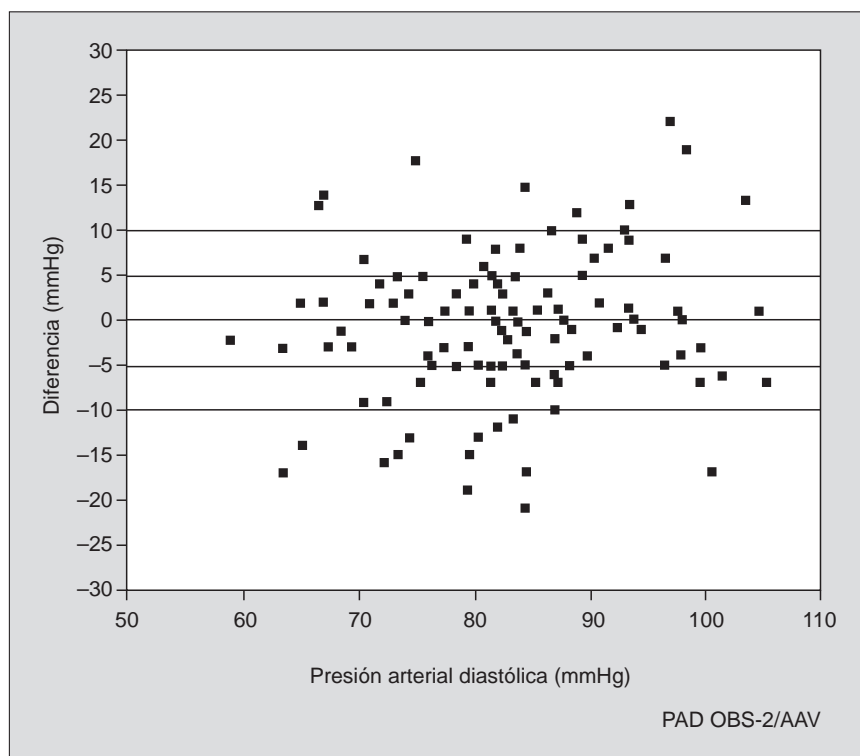


Figura 6. Método de la media de las diferencias individuales (Bland y Altman).

ser influenciado por el azar¹⁰, por lo que no volveremos a hacer referencia a este parámetro.

4. No hemos incluido en el análisis estadístico el coeficiente de correlación de Pearson, dada su inoperancia en este tipo de comparaciones, pues como comentan varios autores 2 medidas pueden estar muy relacionadas (asociación lineal) pero no dar nunca el mismo resultado (concordancia o acuerdo)^{11,24-27}. A pesar de ello, algunas de las recomendaciones vigentes sobre validación de aparatos de medida siguen recomendando este índice como el más apropiado^{15,16,28}, entre otras cosas porque es el más utilizado en recientes publicaciones, argumento que aunque importante por la cantidad no deja de ser irrelevante si en verdad el índice en cuestión no es el adecuado.

Para explicarnos esta «división de opiniones», siguiendo el argot taurino, revisamos la bibliografía ex profeso para encontrar que cada uno de los métodos de estudio de concordancia presenta sus inconvenientes, limitaciones y ventajas.

Así, sobre el método de Bland y Altman se dice que no aporta ninguna información sobre la linealidad de la concordancia entre 2 medidas, y por tanto no valora la magnitud de las diferencias en relación con el valor de la variable^{2,29}. En cambio, otros autores opinan que la representación gráfica puede ser suficientemente ilustrativa para evaluarla¹². Como ya se ha referido con anterioridad, Gómez de la Cámara opina que es el índice ideal para este tipo de variables continuas, ya que utiliza criterios de relevancia clínica que se escaparían con otros métodos, no requiere de una infraestructura estadística compleja y permite cuantificar y ver gráficamente el comportamiento, magnitud y sentido de las diferencias entre procedimientos, siendo un método cada vez más recomendado desde instancias metodológicas^{30,31}. Por otro lado, aunque representa in-

TABLA 3. Coeficiente de correlación intraclass entre los diferentes métodos de medida

| | PAS | IC del 95% | PAD | IC del 95% |
|-----------|------|------------|------|------------|
| OBS1/OBS2 | 0,98 | 1,01-0,94 | 0,92 | 0,99-0,85 |
| OBS1/AAV | 0,81 | 0,92-0,71 | 0,69 | 0,82-0,56 |
| OBS1/OBS3 | 0,90 | 0,97-0,82 | 0,77 | 0,88-0,66 |
| OBS2/AAV | 0,80 | 0,91-0,69 | 0,69 | 0,82-0,56 |
| OBS2/OBS3 | 0,89 | 0,97-0,81 | 0,78 | 0,89-0,67 |
| OBS3/AAV | 0,86 | 0,95-0,76 | 0,81 | 0,91-0,70 |

TABLA 4. Valores de CCI, clasificación de Fermanian

| | |
|------------|---------------------|
| 1 | Máxima concordancia |
| > 0,91 | Muy buena |
| 0,71- 0,90 | Buena |
| 0,51-0,70 | Moderada |
| 0,31-0,50 | Mediocre |
| < 0,30 | Mala concordancia |

intervalos de concordancia de una muestra, al determinar su error estándar y calcular el intervalo de confianza, permite conocer la precisión de dichas estimaciones.

Respecto al CCI, aunque es indudable que proporciona información complementaria al MMD⁴ y, como concluye Prieto¹², es reco-

mendable para cuantificar la fiabilidad, también posee sus limitaciones. Este mismo autor indica que depende de la variabilidad de la muestra; así, si la muestra es heterogénea el CCI tiende a ser mayor y viceversa. Por otro lado, de nuevo el profesor Gómez de la Cámara argumenta que éste es un índice de acuerdo indulgente cuando la variabilidad intrasujetos es inferior a la variabilidad entre sujetos, como es el caso de este estudio, mejorando la apariencia del acuerdo cuando desde el punto de vista clínico la variabilidad de las tomas tensionales sobrepasa el límite establecido como aceptable por los expertos¹⁰.

TABLA 5. Índice de concordancia simple

| IC simple | % ≤ 5 mmHg | % ≤ 10 mmHg |
|-----------|------------|-------------|
| OBS1/OBS2 | | |
| PAS | 86,6 | 96,6 |
| PAD | 84 | 97,4 |
| OBS1/AAV | | |
| PAS | 38,3 | 64,1 |
| PAD | 55,8 | 84,1 |
| OBS1/OBS3 | | |
| PAS | 54,6 | 83,2 |
| PAD | 57 | 89,9 |
| OBS2/AAV | | |
| PAS | 38,6 | 62,1 |
| PAD | 59,7 | 81,5 |
| OBS2/OBS3 | | |
| PAS | 55,9 | 81,3 |
| PAD | 61,9 | 86,4 |
| OBS3/AAV | | |
| PAS | 52,9 | 74,7 |
| PAD | 73,1 | 89,9 |

También habría un sesgo si no se establece a priori la relevancia clínica de las diferencias apreciadas, pues los resultados obtenidos mediante el CCI están expresados en términos absolutos; así, diferencias sistemáticas de 5 unidades para todas las observaciones proporcionarían el mismo CCI tanto si se realizan las mediciones en metros como en kilómetros¹².

Un último inconveniente del CCI es que no se puede obtener con los paquetes estadísticos de uso habitual, por ser de cálculo complejo^{4,12}, motivo que podría explicar su no utilización en algunos estudios de fiabilidad, a pesar de que existen algunas fórmulas simplificadas (sobre todo en el caso de 2 observadores) que facilitarían dicha labor¹².

Un objetivo secundario de este trabajo fue comprobar si la inclusión de un observador teóricamente menos experto (OBS-3) habría de resultar en determinaciones con menor concordancia que los otros observadores. No sólo no fue así, sino que este médico obtuvo los menores IC en todas las comparaciones, por lo que desestimamos como posible fuente de variabilidad la inexperiencia, cuando se parte de la misma base metodológica. Es más, posiblemente la carencia de experiencia en una tarea sea suplida por un interés aún mayor para la interpretación correcta de los resultados o en la aplicación de una técnica; si a ello se añade la falta de algunos «vicios adquiridos» en las tomas de PA, que no pudieran ser corregidos en los demás observadores a pesar del curso previo, podrían explicar las mejores comparaciones obtenidas por dicho observador. La falta de concordancia (MMD) observada en este estudio podría ser atribuida, entre otras posibilidades, a la variabilidad intrínseca de la PA³² y a la utilización de un método indirecto de medida, factores que como se sabe son difícilmente neutralizables. Es más, a

pesar de intentar eliminar los sesgos debidos al observador y al método de medida, estamos convencidos de que su efectividad nunca será completa, ya que la subjetividad de cada observador, por muy entrenado que esté, siempre existirá. Por otro lado la misma deducción se podría utilizar para cualquier otro método de medida y no sólo al MMD.

Aunque una de las medidas que se podría utilizar para disminuir la variabilidad intrínseca de la PA sería la repetición de las tomas en el mismo individuo (técnica de test-retest) para neutralizar esas diferencias individuales por un fenómeno de regresión a la media²⁶, no nos pareció que fuera útil en nuestro estudio, ya que esta medida haría que el tiempo transcurrido para ver a cada paciente fuese mayor entre consultas, por lo que aumentaría la espera para el resto de los pacientes y, paradójicamente, restaría eficacia a esta medida al poder aumentar de forma indirecta la variabilidad intrínseca²⁶.

Hacemos notar como limitaciones de nuestro estudio, en primer lugar, que los valores de CCI siempre inferiores para la PAD, se debe a un artefacto estadístico^{6,12} (el rango de valores que puede adoptar la PAD es menor que el que puede adoptar la PAS). En segundo lugar, que sólo hemos valorado a sujetos hipertensos, dejando de analizar el grado de acuerdo en el rango de valores no considerados como HTA. También podría ser discutible el hecho de que nuestras comparaciones las hemos valorado especialmente frente a un AAV cuando la tendencia es, de forma inversa, la de estandarizar estos aparatos automáticos de medida frente a los EMM. La justificación es que una de las ventajas incuestionables de estos nuevos aparatos de medida es que eliminan el sesgo del observador, factor principal en la valoración de la fiabilidad de los métodos de medida, y por tanto,

correctamente validados, pueden constituir una inestimable fuente de comparabilidad.

Una consideración a tener en cuenta sería la de haber utilizado en nuestro trabajo otros AAV, como podría ser el OMRON HEM 722-C (M4), o el OMRON 732 C-E (MX2), o incluso un aparato semiautomático de automedida de presión arterial caso del HEM 422 C2-E, que además de estar igualmente validados tienen la ventaja de su menor coste, lo que repercutiría en su posible uso de forma generalizada tanto en las consultas de atención primaria como en los propios domicilios de los pacientes, ya que su adquisición sería factible a un mayor número de usuarios^{38,39}.

Numerosas investigaciones han estudiado el componente de la fiabilidad de las observaciones. Sin embargo, un número considerable se centraba en la validación de distintos aparatos automáticos de medida de PA, comparándolos con el método manual, con la finalidad de sustituir el método estándar de medida; por tanto, el objetivo final de estos trabajos era diferente al nuestro, y además casi todos utilizaron procedimientos estadísticos que no son los más aceptados desde el punto de vista metodológico para valorar la fiabilidad de los mismos, como el coeficiente de correlación de Pearson, por lo que sus conclusiones no las hemos tenido en cuenta^{17,24,33-35}. En el estudio PAMELA, de forma similar a nosotros, se documentó que la toma consecutiva entre el random zero y la toma tensional en consulta difería de forma significativa. Aunque en la discusión comentan los autores que probablemente se deba a un sesgo del random zero, esto es cuando menos discutible, no dejando de ser una hipótesis no contrastada³³. En cambio, otros trabajos, a pesar de ser parecidos en algunos de los objetivos intermedios (comparación de medidas

entre diferentes profesionales y métodos), no utilizan métodos estadísticos similares³⁶ o no estudian concordancia³⁷ o la interpretación de alguno de los índices estadísticos difiere de la utilizada por nosotros, por lo que tampoco son comparables⁵.

Si estos resultados han sido obtenidos utilizando todas las medidas a nuestro alcance para evitar las posibles fuentes de error y en condiciones teóricamente ideales, es inevitable plantearnos qué ocurrirá en las condiciones habituales de la consulta diaria.

Es evidente que nuevos estudios con un mayor número de pacientes y con inclusión de un número similar de no hipertensos podrían aclarar las dudas que nos surgen con los resultados de este trabajo.

Entre todas las pruebas diagnósticas utilizadas en medicina clínica, la técnica estándar de Korotkoff para la medición de la PA probablemente sea la única que no ha presentado mejoras tecnológicas sustanciales en los casi 100 años transcurridos desde su descripción, pese a haberse documentado su inexactitud², incluso cuando las determinaciones se realizan en condiciones ideales. Si a ello unimos la universalización progresiva en el uso de los aparatos automáticos de medida de PA, con la validación cada vez mayor de este tipo de instrumentos, y la normativa europea pendiente aún de aprobación que podría condenar para siempre el uso con fines sanitarios del mercurio (dado su toxicidad), nuestra reflexión final sería: esfigmomanómetros de mercurio, método de elección de medida de PA, ¿hasta cuándo?

Agradecimientos

Queremos expresar nuestro agradecimiento a la Unidad Docente de Medicina Familiar y Comunitaria de Córdoba y especialmente a D. Luis Pérula de Torres, técnico de salud de la citada unidad, por su continuado y desinteresado

apoyo, indispensable para la realización de este trabajo.

Asimismo, agradecer a todos nuestros pacientes la colaboración expresa dada para la realización del mismo, pues es evidente que serán siempre la base primordial gracias a los cuales cualquier investigación en medicina, en el área clínica, puede llegar a buen término.

Bibliografía

- Korotkoff NS. On methods of studying blood pressure. Bull Imperial Mil Med Acad St. Petersburg, 1905.
- Pickering TG. Ambulatory monitoring and blood pressure variability. London: Science Press, 1991; 1-14.
- Vinacor F. ¿Cuándo la diabetes es diabetes? JAMA (ed. esp.) 1999; 8: 409-422.
- Jiménez J. Comparación de métodos cuantitativos de medida. FMC 1994; 1: 404-410.
- Montes G, Fernández JA, Prada A, Polonio R, Rodríguez D, Pérula LA. Fiabilidad en la medición de la presión arterial: paciente frente a profesionales de atención primaria. Aten Primaria 2000; 25: 73-77.
- Gómez de la Cámara A, Cruz Martos E, De la Cruz Bértolo J, Landa Goñi J, Guillomía Contreras S, Aurrecoechea R. Análisis de la fiabilidad de tres dosímetros portátiles de glucemia. Comparación de diferentes métodos para el estudio de la fiabilidad de las observaciones clínicas. Med Clin (Barc) 1997; 108: 410-413.
- Bland JM, Altman DG. Statical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1: 307-310.
- Redón J, Miralles A, Lurbe E, Pascual JM, Lozano JV. Excreción urinaria de albúmina durante la noche en la hipertensión arterial esencial. Med Clin (Barc) 1995; 104: 608-611.
- Del Cañizo FJ, De La Cruz FJ, Moreira-Andrés MN, Gómez de la Cámara A, Mayol J, Serrano JM. Grado de acuerdo en la medición de glucemia entre el laboratorio y cuatro glucómetros portátiles. MEDIFAM 1997; 7: 362-266.
- Fleiss JL. The design and analysis of clinical experiments. Nueva York: John Wiley and Son, 1986.
- Alonso J, Prieto L, Antó JM. La versión española del SF-36 Health Survey (Cuestionario de Salud SF-36): un instrumento para la medida de los resultados clínicos. Med Clin (Barc) 1995; 104: 771-776.
- Prieto L, Lamarca R, Casado A. La evaluación de la fiabilidad en las observaciones clínicas: el coeficiente de correlación intraclase. Med Clin (Barc) 1998; 110: 142-145.
- Guidelines Subcommittee, 1999 World Health Organisation-International Society of Hypertension Guidelines for the Management of Hypertension. J Hypertens 1999; 3: 131-132.
- The Joint National Committee of Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. The Sixth Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure. Arch Intern Med 1997; 157: 2443-2446.
- Association for the Advancement of Medical Instrumentation. American National Standard for electronic or automated sphygmomanometers. Washington DC: AAMI, 1987.
- O'Brien E, Petrie J, Littler W, De Sor-te M, Padfield P, O'Malley K et al. The British Hypertension Society protocol for the evaluation of automated and semi-automated blood pressure measuring devices with special reference to ambulatory systems. J Hypertens 1990; 8: 607-619.
- Rose G. Standardisation of observers in blood-pressure measurement. Lancet 1965; 1: 673-674.
- O'Brien E, Mee F, Atkins N, Thomas M. Evaluation of three for self-measurement of blood pressure according to the revised British Hypertension Society Protocol: The Omron HEM-705CP, Philips HP 5332, and Nissei DS-175. Blood Pressure Monitoring 1996; 1: 55-61.
- O'Brien ET, Beevers DG, Marshall HJ. Principios generales de la medición de la tensión arterial. En: ABC de la hipertensión. Barcelona: Ancora, 1996.
- Fermanian J. Mesure de l'accord entre deux juges: cas quantitatif. Rev Epidemiol Santé Publ 1984; 32: 408-413.
- Lee J, Koh D, Ong CN. Statistical methods for assessing observer variability in clinical measures. BMJ 1992; 104: 1491-1494.
- Hernández Aguado I, Porta Serra M, Miralles M, García Benavides F, Bolúmar F. La cuantificación de la variabilidad en las observaciones clínicas. Med Clin (Barc) 1990; 95: 424-429.
- National Committee for Clinical Laboratory Standards. User comparison of quantitative clinical methods using patients samples. Proposed guideline. NCCLS Publication EP9-p. Villanova: NCCLS, 1996.
- Mora-Macià J, Ocón Pujadas J. Validación de un modelo de aparato automático de registro de la presión arterial ambulatoria: el Spacelabs 5200. Med Clin (Barc) 1992; 98: 321-324.
- Rasmussen SL, Torp-Pedersen Ch, Borch-Johnsen K, Ibsen H. Valores normales de presión arterial ambulatoria y diferencias entre presión arterial casual y presión arterial ambulatoria: resultados de una encuesta en población danesa. Ib J Hypertens 1999; 4: 86-95.
- Argimón JM, Jiménez J. Medición de variables. En: Argimon Pallàs JM, Jiménez Villa J, editores. Métodos de investigación aplicados a la atención primaria de salud. Barcelona: Doyma, 1991; 109-121.
- La Calle Remigio. Cartas al director. Aten Primaria 1993; 12: 688.
- Mora-Macià J, Ocón Pujadas J. Aten Primaria 1993; 12: 44.
- Bland JM, Altman BJ. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. Comput Biol Med 1992; 20: 337-340.
- Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. BMJ 1992; 304: 1492-1494.
- McCormack M, Infante-Rivard C, Schick E. Agreement between clinical methods of measurement of urinary frequency and functional bladder capacity. Br J Urol 1992; 69: 17-21.
- Cuesta V, Gómez Moprop MB. Precisión en la toma de tensión arterial: eficacia de los esfigmomanómetros automáticos. Rev Clin Esp 1987; 180: 482-485.
- Mancia G, Sega R, Bravi C, Vito G, Valagussa F, Cesana G et al. Ambulatory blood pressure normality: results from the PAMELA study. J Hypertension 1995; 13: 1337-1390.
- Márquez E, Martín JL. Validación de un monitor automático no invasivo de presión arterial: el ACP-2200. Aten Primaria 1994; 14: 815-819.
- Candela Toha AM. Validación de aparatos y métodos de medida: concordancia sí, correlación no. Med Clin (Barc) 1992; 99: 314.

36. Cabedo VR, Silvestre G, García R, Ripoll J, Hernández I. Validez y fiabilidad de la medición de la tensión arterial en un centro de salud. *Aten Primaria* 1995; 15: 15-20.
37. Martínez MA, Aguirre A, Sánchez M, Nevado A, Laguna I, Torre A et al. Determinación de la presión arterial por médico o enfermera: relación con la presión ambulatoria y la masa del ventrículo izquierdo. *Med Clin (Barc)* 2000; 20: 770-775.
38. Córdoba R, Fuertes MI, Álvarez A, Molina I, Solans R, Melero I. Validación de un monitor automático de presión arterial. *Aten Primaria* 1997; 5: 247-250.
39. Artigao LM, Llavador JJ, Sanchis C, López Abril J, Torres C, Rubio M et al. Validación de aparatos automáticos de presión arterial. *Hipertensión* 1998; 15 (n.º extraordinario): 141-156.